Fall 2002 Genomics Exam #1
Genomic Medicine and Sequencing Tools

There is no time limit on this test, though I have tried to design one that you should be able to complete within 3 hours, except for typing and web searches. There are three pages for this test, including this cover sheet. You are not allowed discuss the test with anyone until all exams are turned in at 11:30 am on Friday September 27. **EXAMS ARE DUE AT CLASS TIME ON FRIDAY SEPTEMBER 27**. You may use a calculator, a ruler, your notes, the book and the internet. However, you are not allowed to obtain and read journal articles as a part of your investigations. These questions are taken from the research literature and I do not want you to simply find the papers and read the answers. This is the Honor Code at its finest.

The **answers to the questions must be typed on a separate sheet of paper** unless the question specifically says to write the answer in the space provided. If you do not write your answers in the appropriate location, I may not find them. You may want to capture screen images as a part of your answers which you may do without seeking permission since your test answers will not be in the public domain.

-3 pts if you do not follow this direction.
Please do not write or type your name on any page other than this cover page.
Staple all your pages (INCLUDING THE TEST PAGES) together when finished with the exam.

Name (please print):

Write out the full pledge and sign:

On my honor I have neither given nor received unauthorized information regarding this work, I have followed and will continue to observe all regulations regarding it, and I am unaware of any violation of the Honor Code by others.

How long did this exam take you to complete (excluding typing)?

20 points

1) ATGAACAAGAAACAAAATTTTTACGCAGCCATTATTGTGGCTATTTTTCTTTGTTTGCAA
TTGTCTCATGGCTCTTCAGGTGTCAGCTTTGAAAAAACCCCTGCTATTAAAATTGTAGGA
AACAAATTCTTTGACTCTGAGAGTGGGGAACAGTTCTTCATCAAGGGCATTGCTTACCAA
TTGCAGAGAAGTGAAGAGGAGCTTAGCAATGCAAATGGGGCTTTTGAGACAAGTTATATT
GATGCCTTAGCGGACCCAAAAATATGCTTAAGAGATATTCCATTTTTGAAAATGCTAGGA
GTGAACACACTGCGTGTTTATGCAATAGATCCGACAAAATCACATGATATATGTATGGAA
GCTCTATCTGCCGAAGGAATGTACGTCCTATTAGATCTTTCAGAGCCTGATATTTCAATA
AATAGAGAAAATCCATCTTGGGATGTACATATTTTTGAAAGGTACAAATCTGTAATTGAT
GCAATGTCATCCTTTCCAAATTTGTTGGGGTATTTTGCAGGGAATGAAGTGACAAATGAC
CATACAAATACCTTTGCATCACCCTTCGTGAAGGCCGCAATCCGAGATGCCAAGGAGTAT
ATTTCGCATTCTAATCATAGAAAAATCCCTGTCGGCTATTCAACTAATGACGATGCTATG

Tell me as much as you can about this sequence. Use as many sites as you want to fill me in on all the scoop. However, to receive maximum points, be sure and tell me every web site/database you visit and what you found there, even if you found nothing. Sometimes that is important information too.

I encourage you to take screen shots of any graphics you find helpful. Copy and paste these into your Word file.

Do not report any information about DNA microarrays for this gene. That will be on the next test.

There were two acceptable answers for this question.
The best answer was that the sequence came from a yeast ORF called YLR343W. This is where I got the sequence.
However, based on the BLAST results, it was also acceptable to say the sequence came from one gene that has many names – specifically PBR1 and FKS1.
I wanted each person to analyze the full gene by CD and Kyte-Doolittle. IN addition, you needed to print out results along the way.
The best answers blended data with the thought processes involved in the determination of the sequence's identity.

15 points

2) This question is a simple one but the answer will take you a while to find. Locate an October 2002 publication by Steen, Lien, Madsen and Birkeland. From this starting point, I want you to show me a picture (screen shot) of which two amino acids are being discussed within the 3D structure from the *E. coli* ortholog. Furthermore, I want you to tell me what type of secondary structure these two amino acids help form in the *E. coli* enzyme. You should paste your screen shot/s into your Word file.

In order to get full credit for this rather challenging question, you **must** print out one page from each intermediate step in the process. I do not want every web site your browser displays; print only the pages that represent major steps in the process you undertook to accomplish this goal.

I expected you to document sequences used.

I expected documentation of how you knew which amino acids were the right two. For example, I used BLAST2 to ling up the two orthologs of IDH.

I expected to see documentation that you had found the correct paper.

Finally, I expected a screen shot similar to this. You should have indicated that these two amino acids are a part of the beta sheet structure, but this depended on how you determined structure. I used QuickPDB.

Ideally, you would have chosen a wt protein such as *wt E. coli* IDH (e.g. 1AI2).
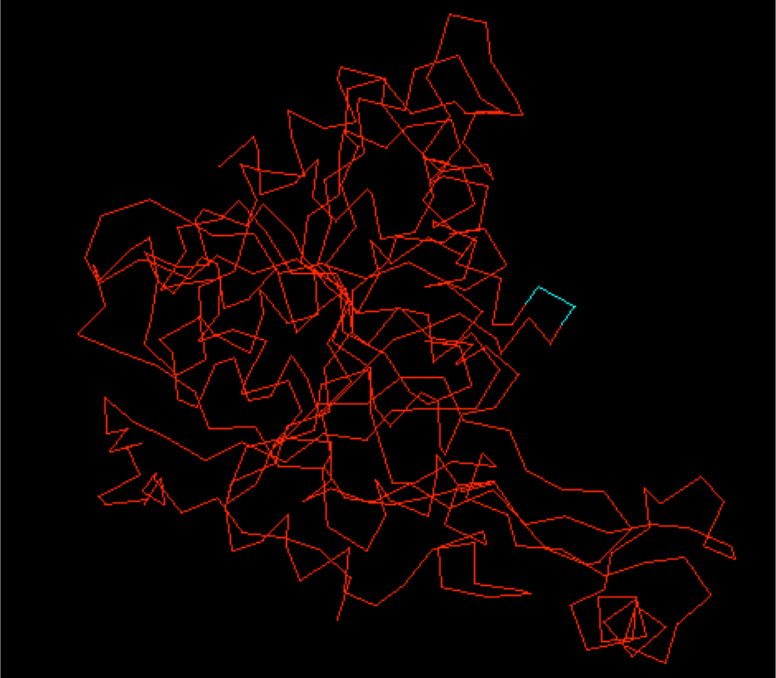
```
Sequence: drag or click to select residues | 3D: double click to select residue

1AI2:_    211   RAAIEYAIANDRDSVTLVHKGNIMKFTEGAFKDWGYQLAREEFGGELIDGGPWLKVKNPNTGKEIVIKDV
1AI2:_    281   IADAFLQQILLRPAEYDVIACMNLNGDYISDALAAQVGGIGIAPGANIGDECALFEATHGTAPKYAGQDK
1AI2:_    351   VNPGSIILSAEMMLRHMGWTEAADLIVKGMEGAINAKTVTYDFERLMDGAKLLKCSEFGDAIIENM

Cursor at 1AI2:_ 391Y

           Polymers:

           1AI2:_

   Stereo

Mouse:
```

20 points

3)

```
CGCTCCGCTGCCTAAGGGCCCCTCGCCACCGCCACCATGGACGCCATCAAGAAGAAGATGCAGATGCTGA
AGCTCGACAAAGAGAACGCCTTGGATCGAGCTGAGCAAGCGGAGGCTGATAAGAAGGCGGCGGAAGACCG
GAGCAAGCAGCTGGAAGATGAGCTGGTGTCACTGCAAAAGAAACTCAAGGGCACTGAAGATGAACTGGAC
```

Tell me everything you can about this sequence. Use as many sites as you want to fill me in on all the scoop. However, to receive maximum points, be sure and tell me every web site/database you visit and what you found there, even if you found nothing. Sometimes that is important information too.

I encourage you to take screen shots of any graphics you find helpful. Copy and paste these into your Word file.

Do not report any information about DNA microarrays for this gene. That will be on the next test.

This sequence came from the mouse α-tropomyosin. You should have determined its cytogenetic location, its cellular role in muscles, OMIM data, orthologs, and PDB structure with a screen shot.

15 points
4) Summarize the main genomic information behind the disease malignant hyperthermia.
I was looking for two main points. First, you had to use OMIM and tell me what this disease is.
Second, you had to list more than one gene that contributes to the disease. At least one gene had to be ryanodine receptor and the type of mutations that can lead to MH.

15 points
5) a) Tell me if the human gene sonic hedgehog is well conserved in fish and mouse. Support your answer with a printout.
Either use Genome Browser or BLAST to show similarity. BLAST 2 would be ideal.
b) What is the cytogenetic position of the human sonic hedgehog gene?
7q36.3
c) What is the cytogenetic position of the mouse sonic hedgehog ortholog?
Chromosome 5, 16 cM down.
d) How many nucleotides are in the human EST for this gene? Support your data with a printout of the EST sequence.
There were several acceptable ETS. Needed to see printout. Full length cDNA/mRNA did not count as an EST.
e) What are the primer sequences and PCR product length for the mouse STS that marks sonic hedgehog?
Again, more than one acceptable answer. Needed to see printout.

15 points
6) Download the May 9, 2000 *PNAS* (*Proceedings of the National Academy of Sciences, USA*) publication: Vol. 97 (10): 5334-5339. Do not use the library copy.

You only need to read the abstract and figure 5 to answer this question, though you are allowed read other parts of the paper if you want. Print the page with figure 5 and do your work on this page. Be sure to turn in your marked up page as a part of your answer.

How many base changes have occurred between the γ DNA Rep Prot (#103) and the α, β, γ IHV DNA pol (#51)?

First, I needed to see a printout. PNAS is freely available, so PDF version was available to you.

All you had to do was measure X axis distances and compare to scale bar.
However, this turned out to be a deceptive problem. Since the scale bar showed single base changes, and there cannot be partial changes (e.g. half a base change), you should have recognized that all X axis distances were in full base change units. Therefore, the only correct answer was 17 bases. No fractions were possible if you understood that all changes were full bases only.